
Journal of Quality Technology
A Quarterly Journal of Methods, Applications and Related Topics

ISSN: 0022-4065 (Print) 2575-6230 (Online) Journal homepage: https://www.tandfonline.com/loi/ujqt20

# Decomposition of *T*2 for Multivariate Control Chart Interpretation

**Robert L. Mason, Nola D. Tracy & John C. Young**

**To cite this article:** Robert L. Mason, Nola D. Tracy & John C. Young (1995) Decomposition of *T*2 for Multivariate Control Chart Interpretation, Journal of Quality Technology, 27:2, 99-108, DOI: 10.1080/00224065.1995.11979573

**To link to this article:** https://doi.org/10.1080/00224065.1995.11979573

# Decomposition of $T^2$ for Multivariate Control Chart Interpretation

ROBERT L. MASON

*Southwest Research Institute, San Antonio, TX 78228-0510*

NOLA D. TRACY and JOHN C. YOUNG

*McNeese State University, Lake Charles, LA 70609-2340*

Multivariate control charts using Hotelling's $T^2$ statistic are popular and easy to use but interpreting their signals can be a problem. Identifying which characteristic or group of characteristics is out of control when the chart signals often necessitates an examination of the univariate charts for each variable. It is shown in this paper that the interpretation of a signal from a $T^2$ statistic is greatly aided if the corresponding value is partitioned into independent parts. Information on which characteristic is significantly contributing to the signal is readily available from this decomposition.

## Introduction

THE use of multivariate control charts to monitor industrial processes is increasingly popular. This is a result of the many recent advances that have occurred in multivariate quality control, such as in multivariate cumulative sum (CUSUM) control charts (e.g., Crosier (1988), Healy (1987), Pignatiello and Runger (1990), and Woodall and Ncube (1985)) and multivariate exponentially weighted moving average (EWMA) control charts (e.g., Lowry, et al. (1992)), as well as the improved effectiveness of these techniques to identify the cause of an out-of-control signal. While it is common in industry to monitor individual process characteristics with separate univariate charts, more attention is being given to procedures that combine multiple characteristics into a single chart.

Among the most popular multivariate control charts is the one based on Hotelling's $T^2$ statistic (see Jackson (1985) or Tracy, Young, and Mason (1992)). Because of its similarity to a standardized univariate Shewhart control chart, it is often called a multivariate Shewhart-type control chart. A major advantage of Hotelling's $T^2$ statistic is that it can be shown to be the optimal test statistic for detecting a general shift in the process mean vector for an individual multivariate observation (Hawkins (1991)). However, the technique has several practical drawbacks. A major drawback is that when the $T^2$ statistic indicates that a process is out of control, it does not provide information on which variable or set of variables is out of control. Further, it is difficult to distinguish location shifts from scale shifts since the $T^2$ statistic is sensitive to both types of process changes.

Several recent papers address these problems and present alternative approaches for such multivariate charts. Doganaksoy, Faltin, and Tucker (1991) propose ranking the components of an observation vector according to their relative contribution to a signal using an univariate $t$ statistic as a criterion. Hawkins (1993), as well as Wade and Woodall (1993), uses regression adjustments for individual variables to improve the diagnostic power of multivariate $T^2$ charts following a chart signal. While these approaches are intuitively appealing and effective in many standard process control situations, there arises the question of whether the techniques can be improved and more closely connected to the $T^2$ control chart statistic.

The purpose of this paper is to present a unified approach to use in identifying the source of signals given by a $T^2$ statistic in a multivariate control

Dr. Mason is Manager of the Statistical Analysis Section. He is a Senior Member of ASQC.

Dr. Tracy is an Associate Professor in the Department of Mathematics, Computer Science, and Statistics. She is a Member of ASQC.

Dr. Young is a Professor in the Department of Mathematics, Computer Science, and Statistics.

chart. It will be shown that the ranking technique of Doganaksoy, Faltin, and Tucker (1991) as well as Hawkins' (1993) regression adjustment are subsets of this procedure, and that these ideas can be extended to cover the range of situations that can arise in using multivariate procedures. Our procedure consists of decomposing the $T^2$ statistic into independent parts, each of which is similar to an individual $T^2$ variate. While this paper focuses on individual observations, a similar technique could be applied to subgroup means with minor modifications to the various formulas.

## Decomposition

Let $\mathbf{X}_i = (X_{i1}, X_{i2}, \ldots, X_{ip})'$ represent a $p$-dimensional vector of measurements made on a process at time period $i$. The value $X_{ij}$ represents an observation on the $j^{\text{th}}$ characteristic. Assume that when the process is in control, the $\mathbf{X}_i$ are independent and follow a multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Some of our results, such as the determination of cutoff values for our proposed statistics, are dependent on this distribution assumption, and it may be necessary to check its validity using a multivariate normal goodness-of-fit test (e.g., Gnanadesikan (1977)).

We estimate $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ from a reference sample having $n$ observations using $\overline{\mathbf{X}}$ and $\mathbf{S}$, the usual sample mean vector and covariance matrix. If the size of the historical sample is large, it is common to assume that these estimates are equal to the true population parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. However, as noted by Tracy, Young, and Mason (1992), that assumption is not necessary in multivariate charting.

To construct a multivariate Shewhart-type control chart for a process, Hotelling's $T^2$ statistic is used as the charting statistic. This statistic has the form

$$T^2 = (\mathbf{X}_i - \overline{\mathbf{X}})'\mathbf{S}^{-1}(\mathbf{X}_i - \overline{\mathbf{X}}) \qquad (1)$$

which, when multiplied by the constant

$$\frac{n(n-p)}{p(n+1)(n-1)}$$

follows an $F$ distribution with $p$ and $n-p$ degrees of freedom. It would be most helpful in this setting to be able to determine the net effect of each of the $p$ variables on the $T^2$ statistic in (1), and the particular factors determining the effect. Such knowledge would aid in interpreting out-of-control signals from the chart.

Jackson (1980, 1991) discusses decomposing the $T^2$ statistic into a sum of $p$ principal components and using these components to help solve this identification problem. When the principal components represent meaningful groupings of variables, the interpretation of out-of-control signals is readily apparent (see Jeffery and Young (1993)). However, in many cases it is difficult or, even impossible, to attach meaning to the principal components, and the characteristics associated with out-of-control signals cannot be determined.

Another approach to solving the signal interpretation problem is to decompose the $T^2$ statistic into independent components, each of which reflects the contribution of an individual variable. For simplicity, we assume that we wish to group the first $p-1$ variables together and to isolate the $p^{\text{th}}$ variable so that $\mathbf{X}_i = (\mathbf{X}_i^{(p-1)\prime}, \mathbf{X}_{ip})'$, where $\mathbf{X}_i^{(p-1)}$ is a $(p-1)$-dimensional measurement vector excluding the $p^{\text{th}}$ variable. Hotelling's $T^2$ can be partitioned into two parts (see Rencher (1993)) as follows:

$$T^2 = T^2_{p-1} + T^2_{p \cdot 1,\ldots,p-1}\,. \qquad (2)$$

The term $T^2_{p-1}$ is Hotelling's $T^2$ statistic using the first $p-1$ variables, and it has the form

$$T^2_{p-1} = \left(\mathbf{X}_i^{(p-1)} - \overline{\mathbf{X}}^{(p-1)}\right)' \mathbf{S}_{\mathbf{XX}}^{-1} \left(\mathbf{X}_i^{(p-1)} - \overline{\mathbf{X}}^{(p-1)}\right)$$

where $\overline{\mathbf{X}}^{(p-1)}$ is the sample mean vector of the $n$ multivariate observations on the first $p-1$ variables and $\mathbf{S}_{\mathbf{XX}}$ is the $(p-1)\times(p-1)$ principal submatrix of $\mathbf{S}$. Although not shown by Rencher in his decomposition, the statistic $T_{p\cdot 1,\ldots,p-1}$ is the $p^{\text{th}}$ component of the vector $\mathbf{X}_i$ adjusted by the estimates of the mean and standard deviation of the conditional distribution of $X_p$ given $X_1, X_2, \ldots, X_{p-1}$. It is given by

$$T_{p\cdot 1,\ldots,p-1} = \frac{X_{ip} - \overline{X}_{p\cdot 1,\ldots,p-1}}{s_{p\cdot 1,\ldots,p-1}} \qquad (3)$$

where

$$\overline{X}_{p\cdot 1,\ldots,p-1} = \overline{X}_p + \mathbf{b}_p' \left(\mathbf{X}_i^{(p-1)} - \overline{\mathbf{X}}^{(p-1)}\right),$$

$\overline{X}_p$ is the sample mean of the $n$ observations on the $p^{\text{th}}$ variable, $\mathbf{b}_p = \mathbf{S}_{\mathbf{XX}}^{-1}\mathbf{s}_{x\mathbf{X}}$ is a $(p-1)$-dimensional vector estimating the regression coefficients of the $p^{\text{th}}$ variable regressed on the first $p-1$ variables,

$$s^2_{p\cdot 1,\ldots,p-1} = s^2_x - \mathbf{s}'_{x\mathbf{X}}\mathbf{S}^{-1}_{\mathbf{XX}}\mathbf{s}_{x\mathbf{X}}$$

and

$$\mathbf{S} = \begin{bmatrix} \mathbf{S_{XX}} & \mathbf{s}_{x\mathbf{X}} \\ \mathbf{s}'_{x\mathbf{X}} & s^2_x \end{bmatrix} .$$

Since the first term in (2) is a Hotelling's $T^2$ statistic on $p-1$ variables, we can also separately partition it into two parts:

$$T^2_{p-1} = T^2_{p-2} + T^2_{p-1\cdot 1,\ldots,p-2} .$$

The first term, $T^2_{p-2}$, is a Hotelling's $T^2$ statistic on, say, the first $p-2$ variables, and the second term, $T^2_{p-1\cdot 1,\ldots,p-2}$ is the square of the $(p-1)^{st}$ variable adjusted by the estimates of the mean and standard deviation of the conditional distribution of $X_{p-1}$ given $X_1, X_2, \ldots, X_{p-2}$. Continuing to iterate and partition in this manner yields the following general decomposition of Hotelling's $T^2$ for $p$ variables:

$$\begin{aligned} T^2 &= T^2_1 + T^2_{2\cdot 1} + T^2_{3\cdot 1,2} + T^2_{4\cdot 1,2,3} + \cdots + T^2_{p\cdot 1,\ldots,p-1} \\ &= T^2_1 + \sum_{j=1}^{p-1} T^2_{j+1\cdot 1,\ldots,j} . \end{aligned} \tag{4}$$

The final $T^2$ value, $T^2_1$, is Hotelling's $T^2$ statistic for the first variable. It reduces to the square of the univariate $t$ statistic for the initial variable:

$$T^2_1 = \frac{\left(X_{i1} - \bar{X}_1\right)^2}{s^2_1} . \tag{5}$$

To illustrate the mechanics of the decomposition procedure, let us apply the techniques to the following chemical industry example. An electrolyzer converts brine into caustic soda and chlorine gas. Three variables commonly measured to monitor the performance of the electrolyzer are the percentage of impurities ($X_1$), the salt-caustic ratio ($X_2$), and a conversion efficiency measure ($X_3$). A historical coded data set of size $n = 16$ yields the following sample mean vector and covariance matrix:

$$\bar{\mathbf{X}} = \begin{bmatrix} 5.700 \\ 17.889 \\ 16.660 \end{bmatrix}$$

and

$$\mathbf{S} = \begin{bmatrix} 0.070 & -0.086 & 0.352 \\ -0.086 & 1.108 & -2.600 \\ 0.352 & -2.600 & 10.115 \end{bmatrix} .$$

Suppose a new observation is taken from the process and is given by $\mathbf{X} = (5.687, 18.508, 17.410)'$. Using equation (1), the $T^2$ statistic associated with this new observation has a value of 1.595. Applying (4), it is possible to decompose this statistic into three components:

$$T^2 = T^2_1 + T^2_{2\cdot 1} + T^2_{3\cdot 1,2} .$$

Equation (5) can be used to determine $T^2_1$:

$$T^2_1 = \frac{(5.687 - 5.700)^2}{0.070} = 0.002 .$$

Similarly, equation (3) can be used to determine $T^2_{2\cdot 1}$ and $T^2_{3\cdot 1,2}$. To compute $T^2_{2\cdot 1}$ we will need $\bar{X}^{(p-1)}_{2\cdot 1}$ and $s^2_{2\cdot 1}$. These are calculated as follows:

$$\bar{X}^{(p-1)}_{2\cdot 1} = 17.889 - 1.224(5.687 - 5.700) = 17.905$$

and

$$s^2_{2\cdot 1} = 1.108 - (-0.086)(0.070)^{-1}(-0.086) = 1.002.$$

Thus

$$T^2_{2\cdot 1} = \frac{(18.508 - 17.905)^2}{1.002} = 0.363 .$$

In a similar fashion, we can compute $\bar{X}_{3\cdot 1,2}$ and $s^2_{3\cdot 1,2}$:

$$\begin{aligned} &\bar{X}_{3\cdot 1,2} = \\ &16.66 + \begin{bmatrix} 2.369 \\ -2.163 \end{bmatrix}' \left( \begin{bmatrix} 5.687 \\ 18.508 \end{bmatrix} - \begin{bmatrix} 5.700 \\ 17.889 \end{bmatrix} \right) \\ &= 15.290 \end{aligned}$$

and

$$\begin{aligned} &s^2_{3\cdot 1,2} = 10.115 - \\ &\begin{bmatrix} 0.352 \\ -2.600 \end{bmatrix}' \begin{bmatrix} 0.070 & -0.086 \\ -0.086 & 1.108 \end{bmatrix}^{-1} \begin{bmatrix} 0.352 \\ -2.600 \end{bmatrix} \\ &= 3.658 \end{aligned}$$

so that

$$T^2_{3\cdot 1,2} = \frac{(17.410 - 15.290)^2}{3.658} = 1.230 .$$

Adding our results we confirm (4):

$$\begin{aligned} T^2 &= 1.595 = 0.002 + 0.363 + 1.230 \\ &= T^2_1 + T^2_{2\cdot 1} + T^2_{3\cdot 1,2} . \end{aligned}$$

There are some interesting properties associated with the decomposition given in (4). Note first that the ordering of the individual conditional terms is not unique. There are in fact $p!$ different partitionings that will yield the same overall $T^2$ statistic. For example, we could begin the partitioning by selecting any one of the $p$ variables. We could then choose any of the $p-1$ remaining variables to condition on the first selected characteristic. Next we could choose any of the remaining $p-2$ variables to condition on the first two selected characteristics. Continuing in this fashion would yield the different sequences of terms composing the overall statistic. To illustrate, suppose $p = 3$. Then the following $3! = 6$ decompositions of the $T^2$ value for an individual observation are possible:

$$\begin{aligned} T^2 &= T_1^2 + T_{2\cdot 1}^2 + T_{3\cdot 1,2}^2 \\ &= T_1^2 + T_{3\cdot 1}^2 + T_{2\cdot 1,3}^2 \\ &= T_2^2 + T_{3\cdot 2}^2 + T_{1\cdot 2,3}^2 \\ &= T_2^2 + T_{1\cdot 2}^2 + T_{3\cdot 1,2}^2 \\ &= T_3^2 + T_{1\cdot 3}^2 + T_{2\cdot 1,3}^2 \\ &= T_3^2 + T_{2\cdot 3}^2 + T_{1\cdot 2,3}^2 \,. \end{aligned}$$

There are $p$ terms in the $T^2$ partition in (4) and, since each one is squared, they are all positive. Thus, each of the terms increases the value of the overall $T^2$ statistic. If the overall statistic yields an out-of-control signal in a multivariate chart, then it should be possible to determine which term, or terms, accounts for the largest increase in the $T^2$ value. The two terms of the series of greatest interest often are the ones containing the unadjusted contribution of a single selected variable, and the term containing the adjusted contribution of one of the variables after adjusting for the other $p-1$ variables.

## Interpretation of Signals

The primary reason for partitioning the $T^2$ statistic in (4) is to obtain information on which variables significantly contribute to an out-of-control signal. The question arises as to how to use the information given in the various terms of the decomposition. It will be useful to recognize that the $p$ terms within a particular decomposition are independent of one another although the terms across the $p!$ decompositions are not all independent. In addition, each of the terms is distributed (under the appropriate null hypothesis) as a constant times an $F$ distribution having 1 and $n-1$ degrees of freedom. The value of this constant is $((n+1)/n)^{1/2}$ (see Tracy, Young, and Mason (1992)). The exact distribution is as follows:

$$T_{j+1\cdot 1,\ldots,j}^2 \sim \frac{n+1}{n} F_{(1,n-1)} \,. \tag{6}$$

(See Appendix A for details and discussion of (6).) Thus, one can compare each term in (4) to a tabled $F$ value, times a constant, to determine if it is significant. This process provides a mechanism for deciding when a term is signaling a problem.

It is now possible to outline a procedure for utilizing this information. Suppose for a given observation that the overall $T^2$ statistic given in (1) is determined to be significant. This indicates that a problem exists but no information is provided about the variables contributing to this problem. Consider next utilizing the decomposition approach given in (4). If we compute the $p$ conditional $T^2$ terms which condition each variable on the remaining $p-1$ variables, we obtain the following set of values:

$$T_{1\cdot 2,\ldots,p}^2, T_{2\cdot 1,3,\ldots,p}^2, \ldots, T_{p\cdot 1,\ldots,p-1}^2 \,. \tag{7}$$

Each of these terms can be compared to the $F$ distribution given in (6) to determine if it is significant.

Similarly, we can compute the $p$ unconditional $T^2$ terms based on squaring a univariate $t$ statistic. This yields the following set of values,

$$T_1^2, T_2^2, \ldots, T_p^2 \,. \tag{8}$$

Each of these terms also can be compared to the $F$ distribution in (6) to determine if it is significant.

There are many other $T^2$ terms that can be computed and examined using the form of the decomposition in (4). In many situations, several of these conditional $T^2$ values will be significant. One quick method for determining the usefulness of exploring these relationships is to compute the differences between the overall $T^2$ value for the given observation and the values of the terms computed in (6). For example, we could compute $T^2 - T_1^2$, $T^2 - T_2^2$, ..., $T^2 - T_p^2$. If any of these differences are significant, using the fact that each follows an $F$ distribution (see Seber (1984)), they indicate that the remaining conditional values in the specific decomposition should be computed and examined. Since the number of conditional terms in the variations of (4) can be large, it is suggested that only the significant ones be recorded for each observation. Accomplishing this task will require the use of a computer algorithm to reduce the required computational work.

## Comparison to Other Methods

Hawkins (1993) presents three different decompositions of the overall $T^2$ statistic to aid in interpreting multivariate control charts. Two of these are related to the decomposition given in (4). In the first approach, Hawkins defines a set of regression-adjusted variables using the vector

$$\mathbf{Z} = \boldsymbol{\Sigma}^{-1}(\mathbf{X} - \boldsymbol{\mu}) \ . \qquad (9)$$

To simplify the comparisons we will replace $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ with their sample estimates, $\bar{\mathbf{X}}$ and $\mathbf{S}$. Using these substitutions and expanding the left-hand side of (9), it is possible to show (see Appendix B) that the $j^{\text{th}}$ component of the vector $\mathbf{Z}$ is given by

$$Z_j = \frac{T_{j \cdot 1,\ldots,j-1,j+1,\ldots,p}}{s_{j \cdot 1,\ldots,j-1,j+1,\ldots,p}} . \qquad (10)$$

Hawkins notes that $Z_j$ is the standardized residual when the $j^{\text{th}}$ variable is regressed on the remaining $p-1$ variables in $\mathbf{X}$. In addition, $Z_j$ is, apart from a scalar based on the inverse of the conditional standard deviation, the square root of one of the conditional $T^2$ components in the decomposition of the overall $T^2$ statistic. Thus, Hawkins' regression-adjusted statistic is a useful diagnostic element in interpreting a $T^2$ signal, as its magnitude directly impacts the size of the $T^2$ statistic. However, as shown in the decomposition in (4), it is only one of several different conditional $T^2$ values.

In a second approach, Hawkins decomposes the $T^2$ statistic using the standardized residuals from the regression of the $j^{\text{th}}$ variable on the first $j-1$ variables. These are defined by

$$\mathbf{Y} = \mathbf{C}(\mathbf{X} - \boldsymbol{\mu}) \qquad (11)$$

where $\mathbf{C}$ is the Cholesky lower triangular root of $\boldsymbol{\Sigma}^{-1}$. In our notation the $j^{\text{th}}$ component of the vector $\mathbf{Y}$ is given by

$$Y_j = \frac{T_{j \cdot 1,\ldots,j-1}}{s_{j \cdot 1,\ldots,j-1}} \ . \qquad (12)$$

Hence, $Y_j$ also is related to a conditional term in the $T^2$ decomposition.

Another method based on the $T^2$ statistic is the step-down procedure of Roy (1958). It assumes there is an *a priori* ordering among the means of the $p$ variables and it sequentially tests subsets using this ordering to determine the sequence. The test statistic has the form

$$F_j = \frac{T_j^2 - T_{j-1}^2}{1 + T_{j-1}^2/(n-1)}$$

where $T_j^2$ represents the unconditional $T^2$ for the first $j$ variables in the chosen group. In the setting of a multivariate control chart, $F_j$ would be the charting statistic.

Using the result in (2), it can be shown that the numerator of $F_j$ is a conditional $T^2$ value:

$$T_j^2 - T_{j-1}^2 = T^2_{j \cdot 1,2,\ldots,j-1} \ .$$

This term, apart from a scalar, is equivalent to the squared standardized residual given in (11). Thus, Roy's approach is closely related to Hawkins' regression adjustment, and therefore to the decomposition of the $T^2$ statistic.

Another proposed partitioning procedure, given by Murphy (1987), uses the overall $T^2$ value and compares it to a $T^2$ value based on a subset of variables. The difference statistic is given by

$$D_j = T^2 - T_j^2$$

where $D_j$ follows an $F$ distribution apart from a constant. Note that no *a priori* ordering is assumed in this method as in Roy's approach. Using the results in (2) and (4) it can be shown that $D_j$ represents a sum of $p-j$ conditional $T^2$ values. An example of one possible representation is given by

$$D_j = \sum_{i=j+1}^{p} T^2_{i \cdot 1,2,\ldots,i-1} \ .$$

Thus, $D_j$ contains a portion of the conditional terms in the decomposition of the $T^2$ statistic.

Another useful $T^2$ diagnostic tool is the univariate ranking procedure proposed by Doganaksoy, Faltin, and Tucker (1991). It is based on use of the $p$ unconditional $T^2$ terms given in (5). While these are part of the decomposition of the $T^2$ statistic, they represent only a few of the many possible terms. We have shown that a $T^2$ value for $p$ variables can be partitioned into $p$ independent components, each of which has the form of a $T^2$ statistic for a single variable. These individual parts are related to the standardized residuals obtained from the regression of a single specified variable against some subset of the remaining variables. While the terms related to the $Z_j$ or $Y_j$, as well as the terms related to univariate $t$ statistics, are very important in the interpretation of the signal, one or more of the other conditional $T^2$ terms could as well contribute to a large signal value for the overall charting statistic. Determining when this happens can aid in the interpretation of the signal and provide other approaches for use in diagnosis.

When using multivariate control charts, it is often difficult to distinguish between location and scale changes. As Hawkins (1991, p.66) states,

> The effects of mean and variance shifts cannot be completely untangled since a shift in variance, even if not accompanied by a shift in mean, alters the average run length (ARL) of location controls.

Nevertheless, the use of the proposed decomposition procedure does aid in identifying a problem variable or set of variables. A large value for the unconditional $T^2$ term in the decomposition indicates that the associated variable is outside its univariate control limits. Significant conditional $T^2$ terms in the decomposition imply that something is wrong with the relationship between a group of the original variables relative to that displayed in the historical sample. This information thus provides a helpful diagnostic tool for interpreting multivariate control charts.

## Examples

A bivariate application of this procedure can be illustrated using a chemical solution example from Jackson (1985) where 15 observations were measured on two variables. Four additional points (labeled A, B, C, and D), included to indicate "abnormal behavior", are contained in Table 1.

Figure 1 contains a plot of the 19 points on a 95% $T^2$ elliptical control region (see Jackson (1991)). Also included are the separate univariate Shewhart control limits for $X_1$ and $X_2$, with $X_1$ along the horizontal axis and $X_2$ along the vertical axis. Together these univariate limits form a square control region. In this example, the process is in control for the first

TABLE 1. Abnormal Observations from Jackson's Data

| Obs. | $X_1$ | $X_2$ |
|---|---|---|
| A | 12.3 | 12.5 |
| B | 7.0 | 7.3 |
| C | 11.0 | 9.0 |
| D | 7.3 | 9.1 |

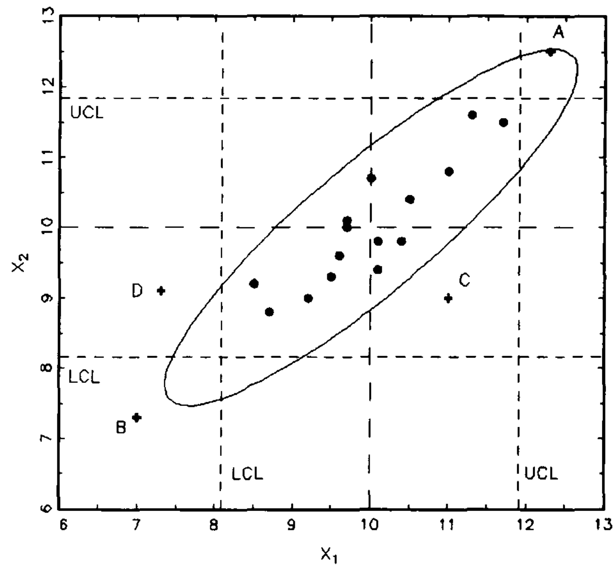


FIGURE 1. Bivariate Control Region for Jackson's Chemical Solution Data.

15 observations, so we can use these observations as historical data to calculate $\overline{\mathbf{X}} = (10, 10)'$ and

$$\mathbf{S} = \begin{bmatrix} 0.799 & 0.680 \\ 0.680 & 0.734 \end{bmatrix}.$$

The $T^2$ statistic signals that the process is out of control for observations B, C, and D (using as an upper control limit $(2 \times 16 \times 14)/(15 \times 13) \times F_{(2,13,0.95)} = 8.743$). This can be seen from the significant $T^2$ values for these points given in the second column of Table 2. The reasons for these out-of-control observations are illustrated in Figure 1. Point B is outside the individual univariate control limits (i.e., the square box) for $X_1$ and $X_2$, but the values are consistent with the relationship between the variables denoted by the ellipse. Point C is outside the ellipti-

TABLE 2. $T^2$ Decomposition for Jackson's Data

| Obs. | $T^2$ | $T_1^2$ | $T_{2\cdot1}^2$ | $T_2^2$ | $T_{1\cdot2}^2$ |
|---|---|---|---|---|---|
| A | 8.51 | 6.62* | 1.89 | 8.51* | 0.001 |
| B | 11.41* | 11.30* | 0.14 | 9.93* | 1.48 |
| C | 23.14* | 1.25 | 21.90* | 1.36 | 21.78* |
| D | 21.59* | 9.13* | 12.50* | 1.10 | 20.49* |

* denotes significance at the 0.05 level, based on one-sided upper critical values: 8.74 (for the overall $T^2$ values) and 4.60 (for the individual components)

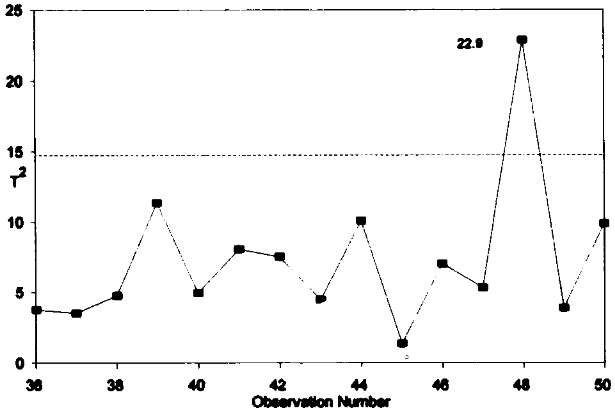


FIGURE 2. $T^2$ Control Chart for Observations 36 through 50 Taken From Hawkins' Data on Switch Drum Dimensions.

cal control region but inside the square box formed by the separate sets of control limits. Individually, the $X_1$ and $X_2$ values are within acceptable ranges but the bivariate observation is not, implying that the relationship between the two variables, possibly caused by a change in the covariance structure, is the problem. Doganaksoy, Faltin, and Tucker (1991) refer to this phenomenon as "counter-correlation", alluding to the apparent inconsistency with the covariance structure. The last point D has its $X_2$ value in an acceptable range, but not its $X_1$ value. Table 2 contains the $T^2$ decomposition values for these four abnormal points. For point B, the conditional values, $T^2_{1\cdot 2}$ and $T^2_{2\cdot 1}$, are small and nonsignificant relative to the remaining portion of the overall $T^2$, and indicate that the relationship between the two variables is appropriate relative to the historical sample. However, the unconditional values, $T_1^2$ and $T_2^2$, are both significantly large, indicating that each variable is outside the univariate Shewhart limits. For point C, the relationship between the variables has changed, as demonstrated by the large and significant conditional $T^2$ values. However, both of the unconditional portions are nonsignificant, reflecting the fact that both the $X_1$ and $X_2$ observations are within their respective control limits. For point D, a change in the relationship between the variables is demonstrated by the significant conditional $T^2$ values but the large unconditional portion for $X_1$ ($T_1^2 = 9.13$) indicates that this variable is the one contributing significantly to the signal.

Another example of the decomposition technique will be illustrated using simulated data from Hawkins (1991) that was based on data taken from Flury and Riedwyl (1988) on five dimensions of switch drums. The variables are the inside diameter of the drum, $X_1$, and the distances from the head to the edges of four sectors cut in the drum, denoted by $X_2$, $X_3$, $X_4$, and $X_5$. A series of 50 observations were sampled and used to represent the process data. After time point 35, an upward shift of 0.25 standard deviations was given to $X_5$, and the marginal standard deviation of $X_1$ was increased by 0.5. No other changes were made to the process variables. Figure 2 contains a $T^2$ chart of the data for observations 36-50 with observations 1-35 serving as the historical data set.

The chart in Figure 2 indicates that the process is out of control at Observation 48. At that instant the value of the overall $T^2$ is 22.88. If we partition this value into independent parts, the values of the two key terms in the decomposition are given in Table 3. Note that the unconditional $T^2$ value for $X_1$ is large and significant at a 0.05 level, as is the conditional $T^2$ value for $X_{1\cdot 2,3,4,5}$. These indicate that the cause for the change in the process is due largely to a combination of a location change in $X_1$ as well as to changes in the relationships between $X_1$ and the remaining variables. While the unconditional $T^2$ value for $X_5$ is not large enough to be significant at a 0.05 level, the conditional $T^2$ value for $X_{5\cdot 1,2,3,4}$ is significant indicating some sort of relationship change between $X_5$ and the remaining variables. The identification of $X_1$ and $X_5$ as the variables contributing significantly to the signal is similar to the conclusions reached by Hawkins when he uses the corresponding $Z_j$ values in a series of univariate Shewhart and CUSUM charts.

Differences between the overall $T^2$ value, 22.88, and the unconditional $T^2$ values in Table 3 are large and indicate that the conditional values should be examined for significance. Some of the difference can be attributed to the conditional $T^2$ term given in the last column of Table 3 while the rest is due to

TABLE 3. Partial $T^2$ Decomposition for Hawkins' Observation 48 ($T^2 = 22.88$)

| Variable ($X_j$) | $T_j^2$ | $T^2_{j\cdot 1,\cdots,j-1,j+1,\cdots,p}$ |
|---|---|---|
| $X_1$ | 7.61* | 16.37* |
| $X_2$ | 0.67 | 0.00 |
| $X_3$ | 0.76 | 1.03 |
| $X_4$ | 0.58 | 1.32 |
| $X_5$ | 3.87 | 6.17* |

* denotes significance at the 0.05 level, based on a one-sided upper critical value of 4.28

TABLE 4. Significant Decomposition Components for Hawkins' Observation 48

| Component | Value | Component | Value | Component | Value |
|---|---|---|---|---|---|
| $T_1^2$ | 7.61 | $T_{1\cdot 2,3,4}^2$ | 15.86 | $T_{5\cdot 4}^2$ | 4.43 |
| $T_{1\cdot 2}^2$ | 8.90 | $T_{1\cdot 2,3,5}^2$ | 16.79 | $T_{5\cdot 1,2}^2$ | 10.48 |
| $T_{1\cdot 3}^2$ | 14.96 | $T_{1\cdot 2,4,5}^2$ | 16.45 | $T_{5\cdot 1,4}^2$ | 11.51 |
| $T_{1\cdot 4}^2$ | 9.41 | $T_{1\cdot 3,4,5}^2$ | 17.41 | $T_{5\cdot 2,4}^2$ | 4.73 |
| $T_{1\cdot 5}^2$ | 15.62 | $T_{1\cdot 2,3,4,5}^2$ | 16.37 | $T_{5\cdot 3,4}^2$ | 4.66 |
| $T_{1\cdot 2,3}^2$ | 15.78 | $T_{3\cdot 1}^2$ | 8.10 | $T_{5\cdot 1,2,3}^2$ | 4.93 |
| $T_{1\cdot 2,4}^2$ | 9.33 | $T_{3\cdot 1,4}^2$ | 6.00 | $T_{5\cdot 1,2,4}^2$ | 11.85 |
| $T_{1\cdot 2,5}^2$ | 15.98 | $T_{3\cdot 1,2}^2$ | 7.06 | $T_{5\cdot 1,3,4}^2$ | 6.89 |
| $T_{1\cdot 3,4}^2$ | 15.18 | $T_{3\cdot 1,2,4}^2$ | 6.70 | $T_{5\cdot 2,3,4}^2$ | 5.66 |
| $T_{1\cdot 3,5}^2$ | 15.10 | $T_{5\cdot 1}^2$ | 11.87 | $T_{5\cdot 1,2,3,4}^2$ | 6.17 |
| $T_{1\cdot 4,5}^2$ | 16.49 | | | | |

the remaining conditional $T^2$ terms. Table 4 contains all the unconditional and conditional $T^2$ values that are significant at a 0.05 level. Each of these 31 elements was compared to a critical value of $(36/35) \times F_{(1,34,0.95)}$, or 4.28. As expected, terms associated with the relationships of $X_1$ and $X_5$ with the remaining variables are significant. In addition, several terms for $X_3$ conditioned on $X_1$ are significant, providing further evidence that there has been a shift in the covariance structure of $X_1$.

The $T^2$ decomposition procedure described above also can be a useful diagnostic tool in the absence of a signal. Doganaksoy, Faltin, and Tucker (1991) describe a situation where counter-correlation is a problem. From a large historical sample containing 1000 observations on three variables they compute $\overline{\mathbf{X}} = (0,0,0)'$, and

$$\mathbf{S} = \begin{bmatrix} 1.0 & 0.7 & 0.6 \\ 0.7 & 1.0 & 0.1 \\ 0.6 & 0.1 & 1.0 \end{bmatrix}.$$

Next, a new observation vector, $\mathbf{X}_{\text{new}} = (2,2,0)'$, is taken and a $T^2$ value of 6.25 is calculated. At the 0.05 significance level, this observation does not signal ($\text{UCL} = (3 \times 1001 \times 999)/(1000 \times 997) \times F_{(3,997,0.95)} = 7.86$). However, when the second variable is ignored, and $T^2$ is recalculated using only the first and third variables, the observation does signal. Doganaksoy, Faltin, and Tucker discuss the difficulty of diagnosing the problem in this case.

Let us re-examine this problem using the proposed decomposition technique. Table 5 shows a complete decomposition of the $T^2$ value, using all three variables, for the new observation. Even though $\mathbf{X}_{\text{new}}$ did not signal in its overall $T^2$ value, it does signal in several components of its $T^2$ decomposition. The significant terms are $T_1^2$, $T_2^2$, $T_{2\cdot 3}^2$, and $T_{1\cdot 3}^2$. The first two significant unconditional terms indicate that $X_1$ and $X_2$ have values outside their acceptable ranges. The two significant conditional terms imply that the relationship of each variable with $X_3$ is a problem.

## Concluding Remarks

Decomposition of a significant $T^2$ value detected in a multivariate control chart can provide a ma-

TABLE 5. $T^2$ Decomposition for Doganaksoy, Faltin and Tucker Data

| Decomposition Number | Decomposition Terms | | |
|---|---|---|---|
| 1 | $T_1^2$ | $T_{2\cdot 1}^2$ | $T_{3\cdot 1,2}^2$ |
| | 4.00* | 0.71 | 1.54 |
| 2 | $T_1^2$ | $T_{3\cdot 1}^2$ | $T_{2\cdot 1,3}^2$ |
| | 4.00* | 2.25 | 0.00 |
| 3 | $T_2^2$ | $T_{3\cdot 2}^2$ | $T_{1\cdot 2,3}^2$ |
| | 4.00* | 0.04 | 2.21 |
| 4 | $T_2^2$ | $T_{1\cdot 2}^2$ | $T_{3\cdot 1,2}^2$ |
| | 4.00* | 0.71 | 1.54 |
| 5 | $T_3^2$ | $T_{2\cdot 3}^2$ | $T_{1\cdot 2,3}^2$ |
| | 0.00 | 4.04* | 2.21 |
| 6 | $T_3^2$ | $T_{1\cdot 3}^2$ | $T_{2\cdot 1,3}^2$ |
| | 0.00 | 6.25* | 0.00 |

* denotes significance at the 0.05 level, based on a one-sided upper critical value of 3.85

jor source of information since the particular variable or group of variables contributing significantly to the signal can be identified. The complete decomposition of the $T^2$ statistic into $p$ independent $T^2$ components is not unique as $p!$ different non-independent partitions are possible. For example, three variables produce six different decompositions, while four variables would yield 24 different partitions. For Hawkins' (1991) data, the five variables would result in 120 distinct decompositions of the $T^2$ value for a given observation. Instead of examining the conditional values for the entire set of $p!$ decompositions, it would be more practical to devise a scheme to output only the significant values. An example of this reduced output is given in Table 4 for a five-variable problem. It is important to note that once a signal has been detected, multiplicity of significance tests is not an issue. Each component in the decomposition can be compared to a critical value as a measure of "largeness" of contribution to the signal rather than for statistical significance.

An advantage in decomposing the $T^2$ value is in its ability to consolidate past research findings on the interpretation of $T^2$ signals. Imbedded in this partitioning are the regression-adjusted variables of Hawkins (1991), the step-down approach of Roy (1958), the $T^2$ differences of Murphy (1987), as well as the standardized $t$ values of Doganaksoy, Faltin and Tucker (1991). In addition, many other terms are included, and each provides information on the cause of the signal. Innovative algorithms for computing these terms is an area of future research.

## Appendix A

For illustration, consider the bivariate case ($p = 2$). For a given vector, $\mathbf{X} = (X_1, X_2)'$, the $T^2$ decomposition is given by

$$T^2 = T_1^2 + T_{2\cdot 1}^2$$

where

$$T_{2\cdot 1} = \frac{X_2 - \bar{X}_{2\cdot 1}}{s_{2\cdot 1}}$$

and $\bar{X}_{2\cdot 1}$ and $s_{2\cdot 1}$ are as defined in (3). Suppose the separate univariate Shewhart control region for $X_1$ has limits $a$ and $b$. For a given $X_1 \in (a, b)$, no signal will be generated by $T^2$ if the corresponding $X_2$ value is within the elliptical control region. Similarly, $T^2$ will signal only if $X_2$ is outside the elliptical control region. This is only possible if $X_2 \in g(X_2|X_1)$; that is, if $X_2$ belongs to the conditional distribution of $X_2$ given $X_1$. To understand this mathematically, consider the following argument. Since $X_2 \sim \mathrm{N}(\mu_2, \sigma_2^2)$ and , $\bar{X}_{2\cdot 1} \sim \mathrm{N}(\mu_{2\cdot 1}, \sigma_{2\cdot 1}^2/n)$, it follows that $X_2 - \bar{X}_{2\cdot 1} \sim \mathrm{N}(\mu_2 - \mu_{2\cdot 1}, \sigma_2^2 + \sigma_{2\cdot 1}^2/n)$. Standardizing, we obtain

$$\frac{(X_2 - \bar{X}_{2\cdot 1}) - (\mu_2 - \mu_{2\cdot 1})}{\sqrt{\sigma_2^2 + \sigma_{2\cdot 1}^2/n}} \sim \mathrm{N}(0,1)$$

Under the null hypothesis $H_0\colon X_2 \in g(X_2|X_1)$, $\mu_2 = \mu_{2\cdot 1}$, and $\sigma_2^2 = \sigma_{2\cdot 1}^2$. Thus we can estimate $\sigma_{2\cdot 1}^2$ with $s_{2\cdot 1}^2$to obtain

$$\frac{X_2 - \bar{X}_{2\cdot 1}}{s_{2\cdot 1}\sqrt{1 + 1/n}} \sim t_{(n-1)}$$

or

$$\frac{X_2 - \bar{X}_{2\cdot 1}}{s_{2\cdot 1}} \sim \sqrt{\frac{n+1}{n}}\, t_{(n-1)}$$

or

$$T_{2\cdot 1}^2 = \left(\frac{X_2 - \bar{X}_{2\cdot 1}}{s_{2\cdot 1}}\right)^2 \sim \left(\frac{n+1}{n}\right) F_{1,n-1}.$$

If the null hypothesis is false, something is wrong with the relationship between $X_1$ and $X_2$. A similar argument can be used to extend this result to the $p$-variable case.

## Appendix B

Replacing $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ with their sample estimates, $\bar{\mathbf{X}}$ and $\mathbf{S}$, equation (9) becomes

$$\mathbf{Z} = \mathbf{S}^{-1}(\mathbf{X} - \bar{\mathbf{X}}).$$

For simplicity, we will only derive $Z_p$ but the other $Z_j$ can be obtained in a similar manner.

Partition $\mathbf{S}$ and $(\mathbf{X} - \bar{\mathbf{X}})$ so that

$$\mathbf{S} = \begin{bmatrix} \mathbf{S_{XX}} & \mathbf{s}_{x\mathbf{X}} \\ \mathbf{s}'_{x\mathbf{X}} & s_x^2 \end{bmatrix}$$

and

$$\mathbf{X} - \bar{\mathbf{X}} = \begin{bmatrix} \mathbf{X}^{(p-1)} - \bar{\mathbf{X}}^{(p-1)} \\ X_p - \bar{X}_p \end{bmatrix}.$$

Then

$$\mathbf{S}^{-1} = \frac{1}{s^2_{x\cdot\mathbf{X}}} \times \begin{bmatrix} s^2_{x\cdot\mathbf{X}}\mathbf{S}^{-1}_{\mathbf{XX}} + \mathbf{S}^{-1}_{\mathbf{XX}}\mathbf{s}'_{x\mathbf{X}}\mathbf{X}^{-1}_{\mathbf{XX}} & -\mathbf{S}^{-1}_{\mathbf{XX}}\mathbf{s}_{x\mathbf{X}} \\ -\mathbf{s}'_{x\mathbf{X}}\mathbf{S}^{-1}_{\mathbf{XX}} & 1 \end{bmatrix}$$

where $s^2_{x\cdot\mathbf{X}} = s^2_{p\cdot 1,\ldots,p-1}$, so that

$$\begin{aligned} Z_p =& \frac{1}{s^2_{x\cdot\mathbf{X}}} \times \left[-\mathbf{s}'_{x\mathbf{X}}\mathbf{S}^{-1}_{\mathbf{XX}}(\mathbf{X}^{(p-1)} - \overline{\mathbf{X}}^{(p-1)}) + X_p - \overline{X}_p\right] \\ =& \frac{1}{s^2_{x\cdot\mathbf{X}}} \times \left[X_p - [\overline{X}_p + \mathbf{b}'(\mathbf{X}^{(p-1)} - \overline{\mathbf{X}}^{(p-1)})]\right] \end{aligned}$$

where $\mathbf{b} = \mathbf{S}^{-1}_{\mathbf{XX}}\mathbf{s}_{x\mathbf{X}}$. Using (3) this reduces to

$$\begin{aligned} Z_p &= \frac{1}{s^2_{x\cdot X}}\left[X_p - \overline{X}_{p\cdot 1,\ldots,p-1}\right] \\ &= \frac{T_{p\cdot 1,\ldots,p-1}}{s_{p\cdot 1,\ldots,p-1}}. \end{aligned}$$

## Acknowledgments

The authors wish to thank Dr. Peter Nelson and the referees for their many helpful comments and suggestions which greatly improved this paper.

## References

CROSIER, R. B. (1988). "Multivariate Generalization of Cumulative Sum Quality-Control Schemes". *Technometrics* 30, pp. 291-303.

DOGANAKSOY, N.; FALTIN, F. W.; and TUCKER, W. T. (1991). "Identification of Out-of-Control Quality Characteristics in a Multivariate Manufacturing Environment". *Communications in Statistics - Theory and Methods* 20, pp. 2775-2790.

FLURY, B. and RIEDWYL, H. (1988). *Multivariate Statistics: A Practical Approach.* Chapman & Hall, London.

GNANADESIKAN, R. (1977). *Methods for Statistical Data Analysis of Multivariate Observations.* John Wiley & Sons, Inc., New York, NY.

HAWKINS, D. M. (1991). "Multivariate Quality Control Based on Regression-Adjusted Variables". *Technometrics* 33, pp. 61-75.

HAWKINS, D. M. (1993). "Regression Adjustment for Variables in Multivariate Quality Control". *Journal of Quality Technology* 25, pp. 170-182.

HEALY, J. D. (1987). "A Note on Multivariate CUSUM Procedures". *Technometrics* 29, pp. 409-412.

JACKSON, J. E. (1980). "Principal Components and Factor Analysis: Part I-Principal Components". *Journal of Quality Technology* 12, pp. 201-213.

JACKSON, J. E. (1985). "Multivariate Quality Control". *Communications in Statistics - Theory and Methods* 14, pp. 2657-2688.

JACKSON, J. E. (1991). *A User's Guide to Principal Components.* John Wiley & Sons, Inc., New York, NY.

JEFFERY, T. C. and YOUNG, J. C. (1993). "Monitoring a Chlorine Production Unit by Multivariate Statistical Procedures". *Proceedings of the Symposia on Chlor-alkali and Chlorate Production and New Mathematical Methods and Computational Methods in Electrochemical Engineering* 9314, The Electrochemical Engineering Society, Inc., pp. 136-147.

LOWRY, C. A.; WOODALL, W. H.; CHAMP, C. W.; and RIGDON, S. E. (1992). "A Multivariate Exponentially Weighted Moving Average Control Chart". *Technometrics* 34, pp. 46-53.

MURPHY, B. J. (1987). "Selecting Out of Control Variables With the $T^2$ Multivariate Quality Control Procedure". *The Statistician* 36, pp. 571-583.

PIGNATIELLO, J. J. and RUNGER, G. C. (1990). "Comparison of Multivariate CUSUM Charts". *Journal of Quality Control* 22, pp. 173-186.

RENCHER, A. C. (1993). "The Contribution of Individual Variables to Hotelling's $T^2$, Wilks' $\Lambda$, and $R^2$" *Biometrics* 49, pp. 479-489.

ROY, J. (1958). "Step-down Procedure in Multivariate Analysis". *Annals of Mathematical Statistics* 29, pp. 1177-1187.

SEBER, G. A. F. (1984). *Multivariate Observations.* John Wiley & Sons, Inc., New York, NY.

TRACY, N. D.; YOUNG, J. C.; and MASON, R. L. (1992). "Multivariate Control Charts for Individual Observations". *Journal of Quality Technology* 24, pp. 88-95.

WADE, M. R. and WOODALL, W. H. (1993). "A Review and Analysis of Cause-Selecting Control Charts". *Journal of Quality Technology* 25, pp. 161-169.

WOODALL, W. H. and NCUBE, M. M. (1985). "Multivariate CUSUM Quality Control Procedures". *Technometrics* 27, pp. 285-292.

Key Words: *Hotelling's $T^2$, Multivariate Control Charts, $T^2$ Control Charts.*